4. First-order logic

In this chapter we finish introducing the notion of first-order logic, and connect this notion to satisfaction and truth in structures.

Let a signature $\sigma = (\mathrm{Fcn}, \mathrm{Rel}, \mathrm{Cn}, \mathrm{ar})$ be given. Now in addition to the variables and the logical symbol $\mathbf{e}$ for equality introduced in the chapter 2, we add some more symbols, assumed to be distinct and different from all the previous symbols:

$\mathbf{n}$, a symbol for negation.

$\mathbf{a}$, a symbol for conjunction ("and").

$\mathbf{x}$, a symbol for existential quantification ("there exists").

So altogether the set Symb of *symbols* for the given signature σ is the set

$$\mathrm{Fcn} \cup \mathrm{Rel} \cup \mathrm{Cn} \cup \{v_i : i < \omega\} \cup \{\mathbf{e}, \mathbf{n}, \mathbf{a}, \mathbf{x}\}.$$

Thus a given first-order language is completely determined by its signature; so we may define a first-order language as just a signature. An *expression* is a finite nonempty sequence of members of Symb. Thus terms and equations are expressions. We now introduce some operations on expressions, so that we will rarely see the actual symbols $\mathbf{e}, \mathbf{n}, \mathbf{a}, \mathbf{x}$:

$\neg\varphi = \langle \mathbf{n} \rangle^\frown \varphi$.

$\varphi \wedge \psi = \langle \mathbf{a} \rangle^\frown \varphi^\frown \psi$.

$\exists v_i \varphi = \langle \mathbf{x}, v_i \rangle^\frown \varphi$.

An *atomic equality formula* is an equation. An *atomic non-equality formula* is an expression of the form $R\sigma_0 \ldots \sigma_{m-1}$ for some relation symbol R of rank m and some terms $\sigma_0, \ldots \sigma_{m-1}$. Note that we are being a little sloppy here, we really mean

$$\langle R \rangle^\frown \sigma_0^\frown \cdots {}^\frown \sigma_{m-1}.$$

The collection of *formulas* is the intersection of all sets A of expressions such that:

Each atomic formula (equality or non-equality) is in A.

If φ is in A, then so is $\neg\varphi$.

If φ and ψ are in A, then so is $\varphi \wedge \psi$,

If φ is in A and $i \in \omega$, then $\exists v_i \varphi$ is also in A.

In analogy to Proposition 2.1 we have:

Proposition 4.1. *(i) No proper initial segment of a formula is a formula.*
(ii) If φ is a formula, then exactly one of the following holds:
(a) φ is an atomic equality formula.
(b) φ is an atomic non-equality formula.
(c) φ is $\neg\psi$ for some formula ψ.
(d) φ is $\psi \wedge \chi$ for some formulas ψ, χ.

(e) φ *is* $\exists v_i\psi$ *for some* $i \in \omega$ *and some formula* ψ.

(iii) If φ *and* ψ *are formulas and* $\neg\varphi = \neg\psi$*, then* $\varphi = \psi$.

(iv) If $\varphi, \psi, \varphi', \psi'$ *are formulas and* $\varphi \wedge \psi = \varphi' \wedge \psi'$*, then* $\varphi = \varphi'$ *and* $\psi = \psi'$.

(v) If $i, j \in \omega$*,* φ, ψ *are formulas, and* $\exists v_i\psi = \exists v_j\psi$*, then* $i = j$ *and* $\varphi = \psi$. □

Now we can define the relationship between first-order logic and structures. First a useful abbreviation: if $a \in {}^{\omega}A$, $i \in \omega$, and $c \in A$, then a^i_c is the member of ${}^{\omega}A$ such that for any $j \in \omega$,

$$(a^i_c)_j = \begin{cases} c & \text{if } j = i, \\ a_j & \text{if } j \neq i. \end{cases}$$

Now suppose that $\overline{A}$ is a structure and $a \in {}^{\omega}A$. We define what it means for a to *satisfy* a formula φ in $\overline{A}$, abbreviated by $\overline{A} \models \varphi[a]$. The definition goes by recursion on φ. Proposition 4.1 is needed to assure the unambiguity of the definition. A rigorous justification of definitions by various sorts of recursion can be found in books on set theory; the intent is clear in our case, though.

$$\begin{aligned}
\overline{A} \models \sigma = \tau[a] \quad &\text{iff} \quad \sigma^{\overline{A}}(a) = \tau^{\overline{A}}(a);\\
\overline{A} \models R\sigma_0 \ldots \sigma_{m-1}[a] \quad &\text{iff} \quad \langle \sigma_0^{\overline{A}}(a), \ldots, \sigma_{m-1}^{\overline{A}}(a)\rangle \in R^{\overline{A}};\\
\overline{A} \models \neg\varphi[a] \quad &\text{iff} \quad \text{not}(\overline{A} \models \varphi[a]);\\
\overline{A} \models (\varphi \wedge \psi)[a] \quad &\text{iff} \quad \overline{A} \models \varphi[a] \text{ and } \overline{A} \models \psi[a]);\\
\overline{A} \models \exists v_i\varphi[a] \quad &\text{iff} \quad \text{there is a } c \in A \text{ such that } \overline{A} \models \varphi[a^i_c].
\end{aligned}$$

We define some more logical notions:

$$\begin{aligned}
\varphi \to \psi &= \neg(\varphi \wedge \neg\psi);\\
\varphi \vee \psi &= \neg(\neg\varphi \wedge \neg\psi);\\
\varphi \leftrightarrow \psi &= (\varphi \to \psi) \wedge (\psi \to \varphi);\\
\forall v_i\varphi &= \neg\exists v_i\neg\varphi.
\end{aligned}$$

Proposition 4.2. *Let* $\overline{A}$ *be a structure and* $a \in {}^{\omega}A$*. Then:*

$$\begin{aligned}
\overline{A} \models (\varphi \to \psi)[a] \quad &\textit{iff} \quad \overline{A} \models \varphi[a] \textit{ implies that } \overline{A} \models \psi[a];\\
\overline{A} \models (\varphi \vee \psi)[a] \quad &\textit{iff} \quad \overline{A} \models \varphi[a] \textit{ or } \overline{A} \models \psi[a] \textit{ (or both)};\\
\overline{A} \models (\varphi \leftrightarrow \psi)[a] \quad &\textit{iff} \quad (\overline{A} \models \varphi[a] \textit{ iff } \overline{A} \models \psi[a]);\\
\overline{A} \models \forall v_i\varphi[a] \quad &\textit{iff} \quad \overline{A} \models \varphi[a^i_c] \textit{ for every } c \in A.
\end{aligned}$$

□

We give some examples of these notions for the signatures introduced at the beginning.

Partial orderings. Consider the partial ordering $\overline{A} \stackrel{\text{def}}{=} (\mathscr{P}(\omega), \subseteq)$ of all subsets of ω.

$\overline{A} \models v_0 < v_1[\emptyset, \{0\}, \{0\}, \ldots]$ since the empty set is a proper subset of any nonempty set.

$\overline{A} \models \forall v_0\forall v_1\exists v_2[(v_0 < v_2 \vee v_0 = v_2) \wedge (v_1 < v_2 \vee v_1 = v_2)[a]$ for any $a \in {}^{\omega}A$ since for any $a_0, a_1 \subseteq \omega$ we can take $a_2 = a_0 \cup a_1$ to satisfy this formula.

Groups. For any group $\overline{A}$ we have:

$\overline{A} \models \forall v_0[v_0 \cdot v_0 = e][a]$ iff every non-identity element of $\overline{A}$ has order 2.

$\overline{A} \models (\neg(v_2 = e) \wedge v_2 \cdot (v_2 \cdot v_2) = e)[a]$ iff a_2 has order 3.

Rings. For any structure $\overline{A}$ for this language we have

$\overline{A} \models \forall v_0(v_0 + (-v_0) = 0)[a]$ iff $a + -a = 0$ for all $a \in A$; this expresses one of the usual axioms for rings.

Ordered fields. For $\overline{A}$ any ordered field, $\forall v_0[\neg(v_0 = 0) \rightarrow 0 < v_0 \cdot v_0][a]$ is the little theorem that the square of any nonzero element is positive.

Having introduced the main relationship between structures and logic, we can now define the most important concepts derived from this relationship. Let $\overline{A}$ be a stucture, φ a formula, **K** a class of similar structures, and Γ a set of formulas.

- φ *holds in* $\overline{A}$ iff $\overline{A} \models \varphi[a]$ for every $a \in {}^{\omega}A$. In this case we also say that φ is *true in* $\overline{A}$, or that $\overline{A}$ is a *model of* φ.

- $\overline{A}$ is a *model of* Γ iff it is a model of each member of Γ.

- $\Gamma \models \varphi$ iff every model of Γ is a model of φ. We read this as "Γ models φ".

- $\mathbf{K} \models \varphi$ iff every member of **K** models φ.

Now in this chapter we go through some of the simplest notions concerning models. We begin by applying the material of Chapter 3 concerning sentential logic. Now a tautology in a first-order language is a formula which can be obtained from a sentential tautology by simultaneously replacing each subformula S_i by φ_i, for some formulas φ_i.

Theorem 4.3. *If φ is a tautology of signature σ, then φ holds in every structure of signature σ.*

Proof. Let φ be obtained from a sentential tautology ρ by simultaneously replacing each subformula S_i by χ_i, for some formulas χ_i, for each $i \in \omega$. (Of course only finitely many S_i's actually occur in ψ.) Let $\overline{A}$ be any structure of signature σ, and let $b \in {}^{\omega}A$; we want to show that $\overline{A} \models \varphi[b]$. To this end we produce a sentential assignment f. Namely, for each $i \in \omega$ let

$$f(i) = \begin{cases} 1 & \text{if } \overline{A} \models \chi_i[b], \\ 0 & \text{otherwise.} \end{cases}$$

Then we claim:

(*) For any subformula ψ of ρ, if ψ' is obtained from ψ by simultaneously replacing each subformula S_i of ψ by χ_i, for each $i \in \omega$, then $\overline{A} \models \psi'[b]$ iff $\psi[f] = 1$.

We prove this by induction on ψ:

If ψ is S_i, then ψ' is χ_i, and our condition holds by definition. If inductively ψ is $\neg\tau$, then ψ' is $\neg\tau'$, and

$$\begin{aligned}\overline{A} \models \psi'[b] \quad &\text{iff} \quad \text{not}(\overline{A} \models \tau'[b])\\ &\text{iff} \quad \text{not}(\tau[f] = 1)\\ &\text{iff} \quad \tau[f] = 0\\ &\text{iff} \quad \psi[f] = 1.\end{aligned}$$

Finally if inductively ψ is $\tau \wedge \xi$, then ψ' is $\tau' \wedge \xi'$, and

$$\begin{aligned}\overline{A} \models \psi'[b] \quad &\text{iff} \quad \overline{A} \models \tau'[b] \text{ and } \overline{A} \models \xi'[b]\\ &\text{iff} \quad \tau[f] = 1 \text{ and } \xi[f] = 1\\ &\text{iff} \quad \psi[f] = 1.\end{aligned}$$

This finishes the proof of (*), and with $\rho = \varphi$ in (*) the theorem is proved. $\square$

Theorem 4.4. (Disjunctive normal form) *Suppose that Γ is a set of formulas. Let Δ be the smallest set of formulas containing Γ and closed under $\neg$ and $\wedge$. Suppose that $\varphi \in \Delta$ and φ has a model. Then there exist a positive integer m, a sequence $\psi \in {}^m\Gamma$, and a set $M \subseteq {}^m2$, such that*

$$\emptyset \models \varphi \leftrightarrow \bigvee_{\varepsilon \in M} \bigwedge_{i<m} \psi_i^{\varepsilon(i)},$$

where χ^1 is χ and χ^0 is $\neg\chi$, for any formula χ.

Proof. For each finite subset F of Γ let Θ_F be the smallest set of formulas containing F and closed under $\neg$ and $\wedge$. Note that

$$\Delta = \bigcup\{\Theta_F : F \text{ is a finite subset of } \Gamma\}.$$

Hence there is a finite subset F of Γ such that $\varphi \in \Theta_F$. Clearly F is nonempty. Let $m = |F|$, and let $\langle \psi_i : i < m \rangle$ enumerate F. Let Ω be the set of all sentential formulas involving only S_i for $i < m$. For each $\theta \in \Omega$ let θ' be obtained from θ by simultaneously replacing each subformula S_i of θ by ψ_i. Then $\{\theta' : \theta \in \Omega\}$ contains F and is closed under $\neg$ and $\wedge$. Hence $\Theta_F \subseteq \{\theta' : \theta \in \Omega\}$. So choose $\theta \in \Omega$ so that $\theta' = \varphi$.

(1) θ is true under some sentential assignment.

In fact, we define $f : \omega \to \{0, 1\}$ as follows. Let $\overline{A}$ be a model of φ, and choose any $a \in {}^\omega A$. Define

$$f(i) = \begin{cases} 1 & \text{if } i < m \text{ and } \overline{A} \models \psi_i[a],\\ 0 & \text{otherwise.}\end{cases}$$

Then we claim

(2) For every sentential formula χ involving only S_i with $i < m$, we have $\chi[f] = 1$ iff $\overline{A} \models \chi'[a]$.

This is clear by induction. It follows that $\theta[f] = 1$, as desired in (1).

Now by Theorem 3.3 choose a nonempty $M \subseteq {}^m2$ such that

$$\theta \leftrightarrow \bigvee_{\varepsilon \in M} \bigwedge_{i<m} S_i^{\varepsilon(i)}$$

is a tautology. Then by Theorem 4.3 we have

$$\emptyset \models \varphi \leftrightarrow \bigvee_{\varepsilon \in M} \bigwedge_{i<m} \psi_i^{\varepsilon(i)}. \qquad \square$$

Now we turn to elementary results involving quantifiers. We first need to talk about free and bound occurrences of variables.

Proposition 4.5. *Suppose that σ is a signature, φ is a σ-formula, and $\exists$ occurs in φ at a certain place: say $\varphi_i = \mathbf{x}$ (remember that formulas are just certain sequences, so we can look at the i-th entry of one). Then there is a unique formula ψ of the form $\exists v_j \chi$, χ a formula, which is a segment of φ beginning at the i-th place.* $\square$

In the situation of this proposition, all occurrences of v_j in the indicated ψ-part of φ are said to be *in the scope of the quantifier $\exists$ appearing at the i-th place in φ*, and each such occurrence is said to be a *bound occurrence* of v_j. Thus the notion of "bound" refers to a particular occurrence, and it is always with respect to a particular formula. An occurrence of a variable in a formula φ is a *free occurrence* if it is not a bound occurrence.

Proposition 4.6. *Suppose that σ is a first-order language, $\overline{A}$ is an σ-structure, φ is a σ-formula, $a, b \in {}^\omega A$, and $a_i = b_i$ for every $i \in \omega$ such that v_i occurs free somewhere in φ. Then $\overline{A} \models \varphi[a]$ iff $\overline{A} \models \varphi[b]$.*

Proof. We proceed by induction on φ. If φ is $\sigma = \tau$ for some terms σ and τ, then to occur free in φ means the same thing as just occurring in σ or τ. So by Proposition 2.2 we have $\sigma^{\overline{A}}(a) = \sigma^{\overline{A}}(b)$, $\tau^{\overline{A}}(a) = \tau^{\overline{A}}(b)$, and so

$$\begin{aligned} \overline{A} \models \varphi[a] \quad &\text{iff} \quad \sigma^{\overline{A}}(a) = \tau^{\overline{A}}(a) \\ &\text{iff} \quad \sigma^{\overline{A}}(b) = \tau^{\overline{A}}(b) \\ &\text{iff} \quad \overline{A} \models \varphi[b], \end{aligned}$$

as desired. The case of an atomic formula starting with a relation symbol is treated similarly. The induction hypotheses with respect to $\neg$ and $\rightarrow$ are easy:

$$\begin{aligned} \overline{A} \models \neg\psi[a] \quad &\text{iff} \quad \text{not}(\overline{A} \models \psi[a]) \\ &\text{iff} \quad \text{not}(\overline{A} \models \psi[b]) \\ &\text{iff} \quad \overline{A} \models \neg\psi[a]; \\ \overline{A} \models (\psi \wedge \chi)[a] \quad &\text{iff} \quad \overline{A} \models \psi[a]) \text{ and } \overline{A} \models \chi[a] \\ &\text{iff} \quad \overline{A} \models \psi[b]) \text{ or } \overline{A} \models \chi[b] \\ &\text{iff} \quad \overline{A} \models (\psi \wedge \chi)[b]. \end{aligned}$$

The induction step involving $\exists$ is more delicate. By symmetry it suffices to go one direction only. So, suppose that $\overline{A} \models \exists v_i\psi[a]$. Then there is a $c \in A$ such that $\overline{A} \models \psi[a^i_c]$. Note that $(a^i_c)_j = (b^i_c)_j$ for every j such that v_j occurs free in ψ. Hence by the inductive hypothesis, $\overline{A} \models \psi[b^i_c]$. Hence $\overline{A} \models \exists v_i\psi[b]$. □

This proposition enables us to use more liberal notation for satisfaction of formulas, just as above for terms. Thus we might write $\varphi(x, y)$ to indicate that we are dealing with a formula whose free variables are among x and y (i.e., all free occurrences of variables in φ are free occurrences of x or y), and then write something like $A \models \varphi[a, b]$ for particular elements $a, b \in A$.

A *sentence* is a formula φ such that no variable occurs free in it. In this case we can write simply $\overline{A} \models \varphi$ or $\overline{A} \models \neg\varphi$. A *theory* is a collection of sentences. A theory Γ is *complete* iff for any sentence φ, either $\Gamma \models \varphi$ or $\Gamma \models \neg\varphi$. This is an important notion which will play a role in much of what follows.

Now we turn to the important but somewhat complicated matter of substituting terms in terms, terms in formulas, etc. The first fact is easily shown by induction (exercise!).

Proposition 4.7. *(i) Suppose that τ is a term and we replace zero or more occurrences of a variable v_i in τ by a term σ, obtaining thereby a sequence ρ. Then ρ is a term.*

(ii) Suppose that φ is a formula, τ is a term, and we replace zero or more occurrences of a variable v_i in φ by τ (except not if the occurrence of v_i is just after the symbol $\exists$), obtaining thereby a sequence ψ. Then ψ is a formula. □

The following basic lemma explains how satisfaction is related to substitutions in a formula.

Lemma 4.8. *Let φ be a formula, v_i a variable, τ a term, and suppose*

(*) no free occurrence of v_i in φ is within the scope of a quantifier on a variable occurring in τ.

Let φ' be obtained from φ by replacing every free occurrence of v_i in φ by τ. Then for any structure $\overline{A}$ and any $a \in {}^\omega A$ we have $\overline{A} \models \varphi'[a]$ iff $\overline{A} \models \varphi\left[a^i_{\tau^{\overline{A}}(a)}\right]$.

Proof. We give this proof in full. The harder cases here are the atomic case and the quantifier case.

First we prove the following related fact involving only terms:

(1) If ρ is a term and ρ' is obtained from ρ by replacing all occurrences of v_i in ρ by τ, then for any $a \in {}^\omega A$ we have $\rho'^{\overline{A}}(a) = \rho^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right)$.

We prove this, of course, by induction on ρ. If $\rho = v_j$ with $j \neq i$, then also $\rho' = v_j$, and a and $a^i_{\tau^{\overline{A}}(a)}$ agree at j, so the conclusion is clear. If $\rho = v_i$, then $\rho' = \tau$, and $\rho'^{\overline{A}}(a) = \tau^{\overline{A}}(a) = \rho^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right)$. If k is an individual constant, then $\rho' = \rho$ and the conclusion is obvious.

For the inductive step, suppose that ρ is $\mathbf{F}\tau_0 \ldots \tau_{m-1}$ and we know (1) for each of the terms $\tau_0, \ldots, \tau_{m-1}$. Note that ρ' is $\mathbf{F}\tau'_0 \ldots \tau'_{m-1}$. Then

$$\begin{aligned}
\rho'^{\overline{A}}(a) &= \mathbf{F}^{\overline{A}}(\tau_0'^{\overline{A}}(a), \ldots, \tau_{m-1}'^{\overline{A}}(a)) \\
&= \mathbf{F}^{\overline{A}}\left(\tau_0^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right), \ldots, \tau_{m-1}^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right)\right) \\
&= \rho^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right),
\end{aligned}$$

as desired. This completes the inductive proof of (1).

Now for formulas, we also proceed by induction on φ. In case φ is $\sigma = \rho$, the formula φ' is $\sigma' = \rho'$ with notation as for (1), and so by (1),

$$\begin{aligned}
\overline{A} \models \varphi'[a] \quad &\text{iff} \quad \sigma'^{\overline{A}}(a) = \rho'^{\overline{A}}(a) \\
&\text{iff} \quad \sigma^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right) = \rho^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right) \\
&\text{iff} \quad \overline{A} \models (\sigma = \rho)\left[a^i_{\tau^{\overline{A}}(a)}\right],
\end{aligned}$$

as desired.

Next, suppose that φ is $\mathbf{R}\sigma_0 \ldots \sigma_{m-1}$. Then φ' is $\mathbf{R}\sigma'_0 \ldots \sigma'_{m-1}$. So

$$\begin{aligned}
\overline{A} \models \varphi'[a] \quad &\text{iff} \quad \langle \sigma_0'^{\overline{A}}(a), \ldots, \sigma_{m-1}'^{\overline{A}}(a)\rangle \in \mathbf{R}^{\overline{A}} \\
&\text{iff} \quad \left\langle \sigma_0^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right), \ldots, \sigma_{m-1}^{\overline{A}}\left(a^i_{\tau^{\overline{A}}(a)}\right)\right\rangle \in \mathbf{R}^{\overline{A}} \\
&\text{iff} \quad \overline{A} \models \varphi\left[a^i_{\tau^{\overline{A}}(a)}\right],
\end{aligned}$$

as desired.

Next, suppose that φ is $\neg\psi$, where we know our result for ψ. Then φ' is $\neg\psi'$, and

$$\begin{aligned}
\overline{A} \models \varphi'[a] \quad &\text{iff} \quad \text{not}\ (\overline{A} \models \psi'[a]) \\
&\text{iff} \quad \text{not}\left(\overline{A} \models \psi\left[a^i_{\tau^{\overline{A}}(a)}\right]\right) \\
&\text{iff} \quad \overline{A} \models \varphi\left[a^i_{\tau^{\overline{A}}(a)}\right],
\end{aligned}$$

as desired.

Next, suppose that φ is $\psi \wedge \chi$, where we know our result for ψ and χ. Then φ' is $\psi' \wedge \chi'$, and

$$\begin{aligned}
\overline{A} \models \varphi'[a] \quad &\text{iff} \quad \overline{A} \models \psi'[a] \text{ and } \overline{A} \models \chi'[a] \\
&\text{iff} \quad \overline{A} \models \psi\left[a^i_{\tau^{\overline{A}}(a)}\right] \text{ and } \overline{A} \models \chi\left[a^i_{\tau^{\overline{A}}(a)}\right] \\
&\text{iff} \quad \overline{A} \models \varphi\left[a^i_{\tau^{\overline{A}}(a)}\right],
\end{aligned}$$

as desired.

Finally, suppose that φ is $\exists v_j \chi$.

Case 1. $j = i$. Then $\varphi' = \varphi$ in this case, since v_i does not occur free anywhere in φ. Hence the desired conclusion follows by 4.6.

Case 2. $j \neq i$.

Subcase 2.1. v_i does not occur free in φ. Then $\varphi' = \varphi$, and the desired equivalence is true by 4.6 again.

Subcase 2.2. v_i occurs free in φ. This is the main case. Note that φ' is $\exists v_j \chi'$, where χ' is obtained from χ by replacing every free occurrence of v_i in χ by τ; moreover, (*) holds for χ as well. And by (*) and the assumption of this case, v_j does not occur in τ. Hence:

(**) For any $x \in A$ we have $\tau^{\overline{A}}(a) = \tau^{\overline{A}}(a^j_x)$, and $(a^i_{\tau^{\overline{A}}(a)})^j_x = (a^j_x)^i_{\tau^{\overline{A}}(a)}$.

Now suppose that $\overline{A} \models \varphi'[a]$. Thus $\overline{A} \models \exists v_j \chi'[a]$. Choose $x \in A$ such that $\overline{A} \models \chi'[a^j_x]$. Then by the inductive hypothesis, $\overline{A} \models \chi[(a^j_x)^i_{\tau^{\overline{A}}(a^j_x)}]$. By the above we then get $\overline{A} \models \chi[(a^j_x)^i_{\tau^{\overline{A}}(a)}]$ and further $\overline{A} \models \chi[(a^i_{\tau^{\overline{A}}(a)})^j_x$. Hence $\overline{A} \models \exists v_j \chi[a^i_{\tau^{\overline{A}}(a)}]$, as desired.

The other direction is similar: suppose that $\overline{A} \models \exists v_j \chi[a^i_{\tau^{\overline{A}}(a)}]$. Choose $x \in A$ such that $\overline{A} \models \chi[(a^i_{\tau^{\overline{A}}(a)})^j_x]$. Then $\overline{A} \models \chi[(a^j_x)^i_{\tau^{\overline{A}}(a)}]$, hence $\overline{A} \models \chi[(a^j_x)^i_{\tau^{\overline{A}}(a^j_x)}]$. So by the inductive hypothesis $\overline{A} \models \chi'[a^j_x]$, hence $\overline{A} \models \exists v_j \chi'[a]$. □

From this lemma we can get some important elementary facts, given in the following theorems. First we need another simple lemma

Lemma 4.9. *Let φ be a formula, and let the sequence ψ be obtained from φ by replacing zero or more bound occurrences of v_i by v_j. Then ψ is a formula.*

Proof. By induction on φ. □

Theorem 4.10. (Change of bound variables) *Suppose that $\mathscr{L}$ is a first-order language and φ is a formula of $\mathscr{L}$. Let v_i and v_j be two variables, and suppose that v_j does not occur in φ. Let ψ be obtained by replacing every bound occurrence of v_i in φ by v_j. Then for any $\mathscr{L}$-structure and any $a \in {}^\omega A$ we have $\overline{A} \models (\varphi \leftrightarrow \psi)[a]$.*

Proof. The proof is by induction on φ, and all steps except the quantifier case are easy. So, suppose that φ is $\exists v_k \varphi'$, and we know the theorem for φ'. Note that $k \neq j$, since v_j does not occur in φ. We now consider two cases.

Case 1. $i \neq k$. Then ψ is $\exists v_k \psi'$, where ψ' is obtained from φ' by replacing all bound occurrences of v_i by v_j. Clearly v_j does not occur in φ', so the inductive assumption gives $\overline{A} \models (\varphi' \leftrightarrow \psi')[a]$. Clearly, then, $\overline{A} \models (\varphi \leftrightarrow \psi)[a]$.

Case 2. $i = k$. Then ψ is $\forall v_j \psi'$, where ψ' is obtained from φ' by replacing all occurrences of v_i, free or bound, by v_j. Let ψ'' be obtained from φ' by replacing all bound occurrences of v_i by v_j. Thus ψ' is obtained from ψ'' by replacing all free occurrences of v_i by v_j.

$$\begin{aligned} \overline{A} \models \varphi[a] \quad &\text{iff} \quad \text{there is a } u \in A \text{ such that } \overline{A} \models \varphi'[a^i_u] \\ &\text{iff} \quad \text{there is a } u \in A \text{ such that } \overline{A} \models \varphi'[a^{i\,j}_{u\,u}]. \end{aligned}$$

Here the second equivalence holds because v_j does not occur in φ, hence not in φ', so that 4.6 applies. Now by the inductive hypothesis, for any $b \in {}^\omega A$ we have $\overline{A} \models (\varphi' \leftrightarrow \psi'')[b]$. Hence by the above equivalences,

$$(*) \qquad \overline{A} \models \varphi[a] \quad \text{iff} \quad \text{there is a } u \in A \text{ such that } \overline{A} \models \psi''[a^{i\ j}_{u\ u}].$$

Now no free occurrence of v_i in ψ'' is within the scope of a quantifier on v_j. So, we can apply 4.8, with $\varphi, \varphi', v_i, \tau, a$ replaced by $\psi'', \psi', v_i, v_j, a^{i\ j}_{u\ u}$ respectively. Note that

$$(a^{i\ j}_{u\ u})^i_{v_j^{\overline{A}}(a^{i\ j}_{u\ u})} = a^{i\ j}_{u\ u};$$

hence we get, from (*),

$$\begin{aligned} \overline{A} \models \varphi[a] \quad &\text{iff} \quad \text{there is a } u \in A \text{ such that } \overline{A} \models \psi'[a^{i\ j}_{u\ u}] \\ &\text{iff} \quad \text{there is a } u \in A \text{ such that } \overline{A} \models \psi'[a^j_u] \quad \text{by 4.6} \\ &\text{iff} \quad \overline{A} \models \psi[a]. \end{aligned}$$

□

Theorem 4.11. (Universal specification) *For any formula φ, any variable v_i, and any term τ such that*

(*) no free occurrence of v_i in φ is within the scope of a quantifier on a variable occurring in τ,

and for any $\mathscr{L}$-structure $\overline{A}$ and any $a \in {}^\omega A$ we have

$$\overline{A} \models (\forall v_i \varphi \to \varphi')[a],$$

where φ' is obtained from φ by replacing each free occurrence of v_i in φ by τ.

Proof. Suppose that $\overline{A} \models \forall v_i \varphi[a]$. Then $\overline{A} \models \varphi[a^i_{\tau^{\overline{A}}(a)}]$, so by Lemma 4.8, $\overline{A} \models \varphi'[a]$. □

The funny condition (*) in this theorem is really necessary, and understanding it helps in seeing the real difference between free and bound occurrences. For example, consider the formula $\varphi \stackrel{\text{def}}{=} \forall y \exists x (y < x)$. Then $(\omega, <) \models \varphi[a]$ for any $a \in {}^\omega\omega$, while if we blindly substitute x for y in $\exists x(y < x)$ we obtain the formula $\exists x(x < x)$ which does not hold under any assignment in $(\omega, <)$. We would like our substitution to lead from "true" formulas to "true" formulas.

We can use these theorems to describe exactly what we mean in going from a formula $\varphi(x, y)$ to another formula $\varphi(\sigma, \tau)$, where σ and τ are terms. Here $\varphi(x, y)$ may have free occurrences of variables other than x and y. Let $u_0, \ldots, u_{m-1}$ list all of the variables that appear in σ or τ. We replace any bound occurrences of any of these in φ by new variables, obtaining ψ. (For definiteness, one should take the first m new variables and do the replacement one-by-one. "new" means not occurring in φ, σ, or τ, and different from x and y.) Then, by definition, $\varphi(\sigma, \tau)$ is the result of replacing x and y in ψ simultaneously by

σ and τ. We illustrate this notation in the statement and proof of the following corollary, which in itself is not very important.

Corollary 4.12. *Suppose that x and y are distinct variables. Then*

$$\models \exists x(x = y \wedge \varphi(x, y)) \leftrightarrow \varphi(y, y).$$

Proof. Here we are using the informal notation just introduced. Thus when we write $\varphi(x, y)$ we just mean φ, but we are indicating variables x and y which may appear in φ. Then $\varphi(y, y)$ means that we first replace all bound occurrences of y in φ by some new variable, obtaining thereby a formula ψ, and then in ψ we replace all free occurrences of x by y, obtaining $\varphi(y, y)$. The idea here is that the change of bound variables does no harm, and fixes things up so that the substitution of y for x does not lead to any unwanted "clashes of bound variables".

Now, down to the proof of the corollary. Say x is v_i and y is v_j. First suppose that $\overline{A} \models \exists x(x = y \wedge \varphi(x, y))[a]$. Accordingly, choose $u \in A$ so that $\overline{A} \models (v_i = v_j \wedge \varphi)[a^i_u]$. Thus $a_j = u$ and $\overline{A} \models \varphi[a^i_u]$. Hence by the change of bound variable theorem, $\overline{A} \models \psi[a^i_u]$. Now since $a_j = u$, we have $v_j^{\overline{A}}(a) = u$ and hence $a^i_u = a^i_{v_j^{\overline{A}}(a)}$. Hence by Lemma 4.8 it follows that $\overline{A} \models \varphi(y, y)[a]$.

Conversely, suppose that $\overline{A} \models \varphi(y, y)[a]$. Then by Lemma 4.8 we get $\overline{A} \models \psi[a^i_{v_j^{\overline{A}}(a)}]$. Clearly also $\overline{A} \models (v_i = v_j)\left[a^i_{v_j^{\overline{A}}(a)}\right]$. So $\overline{A} \models \exists x(x = y \wedge \varphi(x, y))[a]$, as desired. $\square$

Lemma 4.13. *(i) $\models \neg\forall x\varphi \leftrightarrow \exists x(\neg\varphi)$.*
(ii) $\models \neg\exists x\varphi \leftrightarrow \forall x(\neg\varphi)$.
(iii) If x does not occur free in ψ, then $\models (\forall x\phi \wedge \psi) \leftrightarrow \forall x(\phi \wedge \psi)$.
(iv) If x does not occur free in ψ, then $\models (\exists x\varphi \wedge \psi) \leftrightarrow \exists x(\varphi \wedge \psi)$.
(v) $\models \exists x(\varphi \vee \psi) \leftrightarrow \exists x\varphi \vee \exists x\psi$.

Proof. (i) and (ii) are clear. For (iii), suppose that $\overline{A}$ is a structure and $a \in {}^\omega A$. First suppose that $\overline{A} \models (\forall x\varphi \wedge \psi)[a]$. Say x is v_i. To show that $\overline{A} \models \forall x(\varphi \wedge \psi)$, let $u \in A$. Since $\overline{A} \models \forall x\varphi[a]$, it follows that $\overline{A} \models \varphi[a^i_u]$, and hence $\overline{A} \models (\varphi \wedge \psi)[a^i_u]$ by 4.6, as desired. If $\overline{A} \models \forall x(\varphi \wedge \psi)[a]$, then for any $u \in A$ we have $\overline{A} \models (\varphi \wedge \psi)[a^i_u]$, hence $\overline{A} \models \varphi[a^i_u]$ and $\overline{A} \models \psi[a]$ by 4.6. So $\overline{A} \models \forall x\varphi \wedge \psi)[a]$, as desired.

And (iv) is proved similarly. Finally, (v) is easy. $\square$

Of course (iii) and (iv) in 4.13 hold also if the quantifiers are on ψ, x not occurring free in φ; we implicitly assume this version of (iii) and (iv) sometimes.

A formula is said to be in *prenex normal form* if it has the form

$$Q_0x_0 \ldots Q_{m-1}x_{m-1}\varphi,$$

where each Q_i is $\forall$ or $\exists$, and φ is quantifier-free.

Theorem 4.14. *For any formula φ there is a formula ψ in prenex normal form such that $\models \varphi \leftrightarrow \psi$ and the same variables occur free in φ and ψ.*

Proof. We go by induction on φ. The atomic case is trivial, and passage to negation is clear by 4.13(i),(ii). Now take φ of the form $\chi \wedge \theta$. So it suffices to show that the conjunction of two formulas in prenex normal form is equivalent to one formula in prenex normal form. By changing bound variables, we may assume that all quantifiers in either of the two disjuncts are on variables not appearing in the other disjunct. Then 4.13(iii),(iv) give the desired result.

This time the induction step from φ to $\exists v_i \varphi$ is trivial. □

EXERCISES

Exc. 4.1. Prove Proposition 4.7.

Exc. 4.2. Suppose that $\varphi, \psi, \chi, \theta$ are formulas, $\models \chi \leftrightarrow \theta$, and ψ is obtained from φ by replacing one or more occurrences of χ in φ by θ. Show that $\models \varphi \leftrightarrow \psi$. Hint: use induction on φ.

Exc. 4.3. Show that $\models \exists x \forall y \varphi \rightarrow \forall y \exists x \varphi$.

Exc. 4.4. Show that

$$\models \exists x[\varphi \wedge \psi \wedge \exists y(\varphi \wedge \neg\psi)] \rightarrow \exists y(\exists x \varphi \wedge \neg x = y).$$

Exc. 4.5. Show that

$$\models \exists x \varphi \wedge \exists y \psi \wedge \exists z \chi \rightarrow \exists x \exists y \exists z[\exists x(\exists y \chi \wedge \exists z \psi) \wedge \exists y(\exists z \varphi \wedge \exists x \chi) \wedge \exists z(\exists x \psi \wedge \exists y \varphi)].$$

Exc. 4.6. In $(\omega, 0, \mathbb{S})$, show that every singleton $\{m\}$ for $m \in \omega$ is definable, i.e., there is a formula $\varphi_m(x)$ with only x free such that $\{m\} = \{a \in \omega : (\omega, 0, \mathbb{S}) \models \varphi[a]\}$. Here $\mathbb{S}$ is the successor function, which assigns $m+1$ to each natural number m.

Exc. 4.7. Let $\mathscr{L}$ be a relational language. A formula φ is *standard* if every nonequality atomic part of φ has the form $\mathbf{R}v_0 \ldots v_{m-1}$, where $\mathbf{R}$ is m-ary. (Normally any sequence of m variables is allowed.) Show that for every formula φ there is a standard formula ψ such that $\models \varphi \leftrightarrow \psi$.

Exc. 4.8. In the language of rings, write down a single sentence whose models are exactly all rings.

Exc. 4.9. We describe an extension of first-order logic that can be used to make the set theoretical notation $\{a \in A : \varphi\}$ formal (rather than being treated as an abbreviation). Let $\mathscr{L}$ be a first order language, with an individual constant $\mathbf{Z}$ which will play a special role (in set theory, this can be the empty set as introduced in a definition). We define *description terms* and *description formulas* simultaneously:

(a) Any variable or individual constant is a description term.

(b) If $\mathbf{O}$ is an operation symbol of positive rank m and $\tau_0, \ldots, \tau_{m-1}$ are description terms, then $\mathbf{O}\tau_0 \ldots \tau_{m-1}$ is a description term.

(c) If $i < \omega$ and φ is a formula, then $Tv_i\varphi$ is a description term. (This is the *description operator*. $Tv_i\varphi$ should be read "the v_i such that φ, or $\mathbf{Z}$ if there is not a unique v_i such that φ".

(d) If σ and τ are description terms, then $\sigma = \tau$ is an atomic description formula.

(e) If $\mathbf{R}$ is an m-ary relation symbol and $\tau_0, \ldots, \tau_{m-1}$ are description terms, then $\mathbf{R}\tau_0 \ldots \tau_{m-1}$ is an atomic description formula.

(f) If φ and ψ are description formulas and $i < \omega$, then the following are description formulas: $\neg\varphi$, $(\varphi \wedge \psi)$, and $\exists v_i \varphi$.

Next we define the value of description terms, and satisfaction of description formulas in an $\mathscr{L}$-structure simultaneously. Let $\overline{A}$ be an $\mathscr{L}$-structure, and let $a \in {}^{\omega}A$.

(a) $v_i^{\overline{A}} = a_i$.

(b) If τ is the term $\mathbf{F}\sigma_0 \ldots \sigma_{m-1}$, then

$$\tau^{\overline{A}}(a) = \mathbf{F}^{\overline{A}}(\sigma_0^{\overline{A}}(a), \ldots, \sigma_{m-1}^{\overline{A}}(a)).$$

(c) If τ is the term $Tv_i\varphi$, then

$$\tau^{\overline{A}} = \begin{cases} \text{the } x \in A \text{ such that } \overline{A} \models \varphi[a] \text{ and } a_i = x & \text{if there is a unique such } x, \\ \mathbf{Z}^{\overline{A}} & \text{otherwise.} \end{cases}$$

(d) If φ is $\sigma = \tau$, then $\overline{A} \models \varphi[a]$ iff $\sigma^{\overline{A}}(a) = \tau^{\overline{A}}(a)$.

(e) If φ is $\mathbf{R}\sigma_0 \ldots \sigma_{m-1}$, then $\overline{A} \models \varphi[a]$ iff $(\sigma_0^{\overline{A}}(a), \ldots, \sigma_{m-1}^{\overline{A}}(a)) \in r_{\mathbf{R}}$.

(f) $\overline{A} \models \neg\varphi[a]$ iff it is not the case that $\overline{A} \models \varphi[a]$.

(g) $\overline{A} \models (\varphi \wedge \psi)[a]$ iff $\overline{A} \models \varphi[a]$ and $\overline{A} \models \psi[a]$.

(h) $\overline{A} \models \exists v_i \varphi[a]$ iff there is an $x \in A$ such that $\overline{A} \models \varphi[a^i_x]$.

Show that for any description formula φ there is an ordinary formula ψ with the same free variables such that $\models \varphi \leftrightarrow \psi$.

Exc. 4.10. We modify the definition of first-order language by using parentheses. Thus we add two symbols (and) to our logical symbols.

We retain in this context the same definition of terms as before. But we change the definition of formula as follows:

An atomic formula is a sequence of one of the following two sorts: $(\sigma = \tau)$, with σ and τ terms; or $\mathbf{R}\sigma_0 \ldots \sigma_{m-1}$, where $\mathbf{R}$ is a relation symbol of rank m and $\sigma_0, \ldots, \sigma_{m-1}$ are terms. Then we define the collection of formulas to be the intersection of all sets of sequences of symbols such that every atomic formula is in A, and if φ and ψ are in A and $i < \omega$, then each of the following is in A: $(\neg\varphi)$, $(\varphi \wedge \psi)$, $(\exists v_i \varphi)$.

Prove the analog of Proposition 4.1(i), adding an additional condition, that in a formula the number of left parentheses is equal to the number of right parentheses, while in any proper initial segment of a formula, either there are no parentheses, or there are more left parentheses than right ones.

Exc. 4.11. Show that 0 and $\mathbb{S}$ are definable in $(\omega, <)$. That is, there are formulas $\varphi(x)$ and $\psi(x, y)$ with only the indicated free variables such that for all $a \in \omega$, $a = 0$ iff $(\omega, <) \models \varphi[a]\}$, and for all $a, b \in \omega$, $\mathbb{S}a = b$ iff $(\omega, <) \models \psi[a, b]$. Here we are working in the language of orderings.